Income Inequality and Its Socioeconomic Impact: Micro-level Evidence from China

Zhu Yi

Supervisor: Professor Li Hongbin

A Thesis Submitted in Partial Fulfillment
of the Requirement for the Degree of
Master of Philosophy
in
Economics

June 2004

Income Inequality and Its Socioeconomic Impact: Micro-level Evidence from China

Abstract

This thesis explores several aspects of income inequality using survey data from China. We first examine the pattern and determinants of income inequality. We show that within-region inequality contributes more than inter-region one to the total inequality. Examining the provinces separately, we find that over half of the provincial inequality is driven by within-village inequality. Although we show that the distribution of farming land is positively associated with village inequality, our cross-section analysis do not find consistent evidence supporting any link between inequality and village-level market development or non-agricultural institutions.

This thesis also tests whether individual health is associated with income and community-level income inequality in China. Although poor health and high inequality are key features of many developing countries, most of the earlier literature has drawn on data from developed countries in studying the association between the two. We find that self-reported health status increases with per capita income, but at a decreasing rate. Controlling for per capita income, we find an "inverted-U" association between self-reported health status and income inequality, which suggests that high inequality in a community poses threats to health. We also find that high inequality increases the probability of health-compromising behaviors such as smoking and alcohol consumption. Most of our findings are robust to different measures of health status and income inequality.

Income Inequality and Its Socioeconomic Impact: Micro-level Evidence from China

摘要

本篇论文采用中国的调查数据研究有关收入不均等的几个课题。我们首先研究了收入不均等的构成状况和决定因素。本篇论文的结果表明在总体收入不均等的构成中，地区内的不均等所占的比重大于地区间的不均等。如果按省来分析，省一级的收入不均等中至少有一半来自村内的不均等。我们进一步研究了影响村内收入不均等的潜在因素。结果显示耕地的不均等对收入不均等有显著正的影响，但是村级的市场发育水平和非农业体制发达程度对于收入不均等没有显著影响。

本篇论文还研究了收入水平和收入不均等与个人健康之间的关系。收入分配不均和健康水平低下是许多发展中国家面临的严重问题，但过去的文献多采用发达国家的数据来研究两者之间的关系。本篇论文的结果显示，人均收入的提高有利于健康的改善，但此种促进作用随着收入的提高而减弱。在人均收入相同的情况下，个人健康与收入不均等呈“倒 U 型曲线”的关系。这说明在收入不均等严重的地区，不均等扩大会导致个人健康恶化。我们的结果还显示，收入不均等的扩大令人更有可能进行一些有损健康的行为，比如吸烟和饮酒。即使我们采用不同的指标来衡量健康和收入不均等，结果是一致的。

Acknowledgement

I would like to gratefully acknowledge contribution and support from my supervisor, Professor Li Hongbin. His expert guidance and valuable help made it possible for me to complete this thesis.

We are very grateful to Professor Loren Brandt, Professor Du Julan, Professor Tsui Kai Yuen and Professor Zhang Junsen for their helpful comments. We also thank the Carolina Population Center for kindly providing the data.

Contents

1	Introduction	P. 1
2	Income Inequality and Its Determinants in China	P. 6
	2.1 Data	P. 6
	2.2 Source and Distribution of Income	P. 7
	2.3 Decomposition of Income Inequality by Geography	P. 8
	2.4 The Determinants of Income Inequality	P. 10
3	Income, Income Inequality and Health in China	P. 12
	3.1 Hypotheses and Previous Literature	P. 12
	3.2 Data and Measurement Issues	P. 18
	3.2.1 Health Indicators	P. 18
	3.2.2 Income Inequality and Relative Income Measures	P. 20
	3.2.3 Other Explanatory Variables	P. 21
	3.3 Estimation Results	P. 22
	3.3.1 Income, Income Inequality and Individual Health	P. 22
	3.3.2 Relative Income and Individual Health	P. 26
4	Conclusion	P. 28
Tables		P. 31
References		P. 52

Tables

Table 1	Source and Distribution of Household Income in Rural China	P. 31
Table 2	Total Inequality and Its Decomposition in Rural China	P. 32
Table 3	Provincial Inequality and Its Decomposition in Rural China	P. 33
Table 4	Descriptive Statistics of Village Variables in Rural China	P. 34
Table 5	OLS Regressions Measuring the Effects of Education and Land on Village Inequality	P. 35
Table 6	OLS Regressions Measuring the Effects of Market Development on Village Inequality	P. 36
Table 7	OLS Regressions Measuring the Effects of Education and Market Development on Village Inequality	P. 37
Table 8	Definitions of Health, Inequality and Other Variables in China	P. 39
Table 9	Descriptive Statistics of Health, Inequality and Other Variables in China	P. 40
Table 10	Correlations between SRHS, PF and ADL in China	P. 41
Table 11	Descriptive Statistics of Variables for Healthy versus Unhealthy People in China	P. 42
Table 12	Probit Regressions Measuring the Effects of Income Inequality on Self-Reported Health Status	P. 43
Table 13	Probit Regressions Measuring the Effects of Income Inequality on Physical Functions	P. 45
Table 14	Probit Regressions Measuring the Effects of Income Inequality on Activities of Daily Living	P. 47
Table 15	Probit and OLS Regressions Measuring the Effects of Income Inequality on Health Behaviors	P. 49
Table 16	Probit Regressions Measuring the Effects of Relative Deprivation on Self-Reported Health Status	P. 50

1 Introduction

The reforms of China's rural sector since the late 1970s are characterized by several main aspects. First, the Household Responsibility System (HRS) replaces former collective agriculture (Lin, 1992). Under HRS, the village distributes land-use rights to households based on household size, and households can keep residual income from agricultural production after meeting tax obligatory and sales quotas to the government procurement. Second, price and marketing reforms permit new entry into input and output markets, and cause decline in government planning of sales in the farm sector, thus increasing the agricultural return and stimulating the output growth (Sicular, 1995; Park and Rozelle, 1998). Third, fiscal decentralization and the policy intending to absorb surplus rural labor into industry lead to the fast development and growth of township and village enterprises (TVEs) during the 1980s (Che and Qian, 1998; Chen and Rozelle, 1999). Finally, relaxation of restrictions on mobility and migration from rural to urban areas brings rural households more access to non-agricultural business (Zhao, 1999).

China has recorded impressive growth over the past 25 years since the introduction of the market economy and implementation of these reforms, and there has been a substantial increase in average living standards. However, in recent years there has been growing concern about the large increase in income inequality during the same period. For example, Bramall (2001) shows that the Gini Coefficient for rural China has increased by almost 50 percent from 1980 to 1999. The rising inequality has had and will have important impacts on various aspects of social life, resulting, for example, in frequent social conflicts (Alesina and Perotti, 1996), higher levels of violent crime (Hsieh and Pugh, 1993), and ultimately in a slowing down of economic growth (Aghion et al., 1999).

The goal of this thesis is to explore several dimensions of income inequality in China, using the high quality household survey data from the China Health and Nutrition Survey (CHNS). We employ regression analysis to examine the pattern of income inequality, the potential factors that may determine income inequality, and a particular aspect of the socioeconomic effects of inequality, i.e., its impact on health.

Most of the previous literature on inequality in China highlights the income disparities across provinces or regions (such as inland versus coastal, and rural versus urban).[1] It is true that economic reform and income growth have taken place unevenly in China, as some provinces (especially in eastern coastal regions) are better situated to take advantage of market liberalization and industrialization opportunities. However, since the reform has broken the pre-reform principle of absolute equality, we might also expect inequality to increase within a single locality. Some recent work shows that the contribution of within-locality inequality is also significant (Gustafsson and Li, 2002; Benjamin et al. 2002; Benjamin et al., 2003).

In order to find a link between the reforms and income inequality in rural China, several studies have emphasized the role of the emerging non-agricultural sector in explaining inequality (e.g., Rozelle, 1994; Khan and Riskin, 1998; Tsui, 1998; Morduch and Sicular, 2002). These papers typically decompose income inequality by source and find that non-farm income is more disequalizing than income from farming. Therefore they suggest that the rising rural inequality is probably driven by the growth of TVEs and increase of non-agricultural opportunities. However, these studies mostly focus on inter-region inequality, and sometimes ignore the inequality within local units. Moreover, the decomposition approach is purely descriptive and does not have a causal interpretation.

There is little evidence on how within-region (province, county or village) inequality is determined. Using regression-based analysis, a recent study by Benjamin et al. (2002) shows exploratory results on the association between human capital, market development and village-level inequality. They find important interactions between education and market variables in determining the village inequality, although their study has a small sample with limited geographic coverage (only 30 villages in two provinces). In another descriptive paper, Benjamin et al. (2003) outline two factors that may affect within-village inequality: the distribution of household endowments, and village-level non-agricultural institutions and market development.

[1] For a review of the literature on spatial inequality in China, see Kanbur and Zhang (1999), and Gustafsson and Li (2002).

The relationship among income, income inequality and health is an issue which has attracted the attention of a variety of social science disciplines such as economics, sociology and public health. From an early stage in the debate, it was argued that income has a positive effect on health (Grossman, 1972; Preston, 1975). This is called the *absolute income hypothesis*. However, some researchers assert that relative income or income inequality plays an equally important role in determining health. According to the *relative income hypothesis* (or the *weak income inequality hypothesis*), people who feel more economically disadvantaged than their peers in a reference group are more likely to have poorer health (Marmot et al., 1991; Wilkinson, 1997). Low relative income may cause stress and depression leading to illness (Cohen et al., 1997) or weaken one's power in the allocation of local health-related resources (Deaton, 2003). Some (Wilkinson, 1996) go even further and argue that income inequality may affect the health of both the poor and the well off in a society (referred to as the *strong income inequality hypothesis*), possibly through disinvestment in public health and human capital, the erosion of social capital, or stressful social comparisons (Kawachi and Kennedy, 1999).

The relative income or income inequality hypotheses has been empirically tested, but almost exclusively drawing on data from industrialized countries, and the results have been mixed.[2] The tests have been conducted at both the aggregate and individual levels. At the aggregate level, a number of studies have shown a robust association between income inequality and public health (e.g., Waldmann, 1992; Kaplan et al., 1996; Kawachi et al., 1997; Lynch et al., 1998). However, the use of aggregate data may be unconvincing. As noted by Gravelle (1998), income inequality may be spuriously correlated with the aggregate measure of health if individual health is a concave function of income. It is therefore difficult to discriminate between the effects of income and income inequality using aggregate data. To differentiate between the absolute income and income inequality effects, recent studies employ individual data. Among these studies, some support the income inequality hypothesis (e.g., Kennedy et al., 1998; Soobader and LeClere, 1999; Blakely et al., 2001), while others find no significant effects of inequality (e.g., Meara, 1999; Blakely et al., 2002; Mellor and

[2]For a systematic review of previous empirical work, see Deaton (2003) and Lynch et al. (2004).

Milyo, 2002).

In this thesis, we use data of a high-quality household survey to explore above dimensions of income inequality in China. We first show that within-region inequality contributes more than inter-region one to the total inequality. Examining the provinces separately, we find that at least one half of the provincial inequality is attributed to within-village inequality. We then use OLS regressions to explore the potential determinants of within-village inequality. Although we show that the distribution of farming land is positively associated with village inequality, our results do not confirm any link between inequality and village-level market development or non-agricultural institutions.

In this thesis we also investigate the relationship between income, income inequality and health in China. We find evidence supporting the absolute income hypothesis, that income has a positive effect on self-reported health status. Consistent with findings by Daly et al. (1998), we also find evidence supporting the strong version of the income inequality hypothesis but not the weak version. However, unlike previous findings of a linear relationship, our results show an "inverted-U" association between self-reported health status and inequality, i.e., the detrimental effect of income inequality on health only appears in communities with high inequality. We also test the effect of relative deprivation and income rank on health and find that only income rank has a significant positive effect on health. This is in contrast with Eibner and Evans (2001), who find relative deprivation more important than rank in explaining individual health. Finally, we also show that rising inequality can significantly increase one's probability of engaging in health-compromising behaviors such as smoking and alcohol abuse.

We contribute to the literature studying the relationship between income inequality and health in the following ways. First, this thesis is one of the first studies to use individual data from a developing country. Although poor health and high inequality are key features of many developing countries, the earlier literature has studied their association drawing mainly on data from U.S. and other industrialized countries.[3] Moreover, as pointed out by

[3]For example, Osler et al. (2002), Shibuya et al. (2002), and Gerdtham and Johannesson (2004) employ data from Japan, Denmark and Sweden, respectively.

Gerdtham and Johannesson (2004), industrial countries like Sweden may not be the best places for studying the effects of income inequality, because these countries are typically more egalitarian and do not have sufficient variation in income inequality across regions. In contrast, China has both rising inequality and a large variation in inequality across localities (Gustafsson and Li, 2002). Second, we extend the previous work by explicitly distinguishing between the relative income hypothesis and the income inequality hypothesis in the same study. Previous studies have tested either the relative income hypothesis (Deaton, 2001; Eibner and Evans, 2001) or the income inequality hypothesis (eg., Mellor and Milyo, 2002).[4] Finally, we measure the income inequality at the community-level, so that our focus is more locally defined than most previous studies, which focus on the state or county level. Using community-level inequality not only facilitates the empirical test by allowing us to work with a larger variation in inequality, but also permits us to examine the potential impacts of inequality within a society by taking a set of people who are more closely related.

The structure of this thesis is as follows. Section 2 discusses the pattern and determinants of income inequality. Section 3 focuses on the relationship between income, income inequality and health. Section 4 concludes.

[4]Gerdtham and Johannesson (2004) test both hypotheses, but their measure of relative income is a simple one.

2 Income Inequality and Its Determinants in China

In this section, we first examine the pattern of income inequality in rural China, and then explore several potential determinants of inequality. The structure of the section is as follows. Subsection 2.1 describes the data. Subsection 2.2 shows the source and distribution of income. Subsection 2.3 decomposes income inequality by geography. Subsection 2.4 reports our regression analysis of the determinants.

2.1 Data

In this thesis, we use the China Health and Nutrition Survey (CHNS) data, which were collected by the Carolina Population Center (CPC) at the University of North Carolina at Chapel Hill, the Institute of Nutrition and Food Hygiene, and the Chinese Academy of Preventive Medicine in 1993.[5] The sample households were randomly drawn in eight provinces including Liaoning, Shandong, Jiangsu, Henan, Hubei, Hunan, Guangxi, and Guizhou. Two cities and four counties were sampled in each province. Four neighborhoods (including suburban villages) in each city, and one county-town neighborhood and three villages in each county, were then randomly selected. Approximately 20 households were sampled per neighborhood or village.

The CHNS data contain detailed information on levels and sources of household income. Specifically, total income is calculated as the sum of income from agriculture such as gardening, farming, raising livestock or poultry, and fishing, and income from non-agricultural business such as wage, family-run business, and welfare subsidies. The data also record a set of village variables which are obtained from a separate community survey. Since we focus on the income inequality of rural China, we restrict our sample to suburban and county villages and exclude those households with non-positive total incomes. In sum, we have 122 villages and 2,276 households in the sample.

We use two inequality indices in our analysis, namely, the Gini coefficient and the Theil

[5] A detailed description of the data and quality control procedures can be obtained at http://www.cpc.unc.edu/projects/china/.

index. The Gini coefficient is most widely used in previous literature, and the Theil index is more helpful in decomposing inequality by geography. We calculate the inequality based on per capita income weighted by the family size. To eliminate the regional differences in price levels, we use the consumer price index included in the CHNS data to adjust all incomes to prices in urban areas in Liaoning province. Basically, the Gini coefficient and the Theil index are highly correlated (0.95 in our sample).

2.2 Source and Distribution of Income

Table 1 shows descriptive statistics for the various sorts of income earned by households in our sample. For each source of income, we report the unconditional means of all values, the proportion and means of only positive values, as well as the Gini coefficient calculated over the observations with positive incomes. We also show the statistics for a smaller sample after excluding suburban villages.

Agriculture is an important economic activity for most rural households, as 50 to 60 percent of the households report positive incomes in each source of gardening, farming and raising livestock or poultry. Jointly, 98 percent of households report positive incomes from at least one source of agriculture. Non-agricultural incomes mainly come from two sources: wage income and an "other" category, which includes such sources as rental income, remittances, pensions and gifts. The proportion of reporting positive incomes from these two sources are both close to 40 percent. Overall, the average household income and per capita income in our sample are 4,758 and 1,187 RMB, respectively. Agricultural income accounts for nearly one third of total income, and wage income contributes most to non-agricultural income.

The last column presents the distribution of each source of income, measured by the Gini coefficient. The Gini for household income for the entire sample is 0.47, and marginally lower for per capita income. Among the agricultural incomes, income from farming is most equitably distributed, possibly due to a fairly egalitarian distribution of land in rural villages. On the other hand, wage income has the most equal distribution among the sources of non-

agricultural income.

After we exclude the suburban village observations, we find that most of the income levels decrease compared with the full sample, as shown in Table 1. This suggests that on average, households in suburban villages are richer than those in county villages. Another pattern of county villages is that the proportion of households reporting positive agricultural incomes is higher, and the proportion reporting positive non-agricultural incomes is lower. The results are natural as households in suburban villages are more likely to engage in non-agricultural business and earn higher incomes. Finally, the distribution of various incomes is very similar to that for the whole sample, although the Gini coefficients are slightly lower.

2.3 Decomposition of Income Inequality by Geography

In this subsection we use Theil index, an additively decomposable inequality measure, to explore how much of the inequality in our sample is at the local level, as opposed to gap in average incomes across regions. We decompose the Theil index in three location levels: province, county and village. Following Gustafsson and Li (2002), we calculate the Theil index as

$$T(y) = \frac{1}{n}\sum_{i}^{n}\left(\frac{y_i}{\mu}\right)log\left(\frac{y_i}{\mu}\right), \tag{1}$$

where y_i is the income of ith observation, μ is the mean income and n is the sample size.

If the sample is divided into m groups, the Theil index can be decomposed as

$$T(y) = \sum_{j}^{m}\left(\frac{n_j}{n}\right)\left(\frac{\mu_j}{\mu}\right)T_j + T(\mu_1 e_1, ..., \mu_m e_m), \tag{2}$$

where n_j and μ_j are the size and mean income of jth group. T_j is inequality within the jth group and e_j is the n_j vector of ones. The two parts in (2) stands for within-group inequality and inter-group inequality respectively, and their sum equals the total inequality in (1).

Table 2 shows the results of our decomposition over the full sample and the sample without suburban villages. Our results indicate that within-location inequality contributes more than inter-location inequality to the total rural inequality, whichever location level is chosen. We report the overall inequality (also measured by the Gini coefficient) and within-

region inequality at different levels. As for the whole sample, 93 percent (column 2) of the total inequality comes from inequality within provinces. The proportion declines as we decompose by smaller geographic units (thus dividing the sample into more groups), but it still exceeds one half at county and village levels (columns 3 and 4). Nearly 60 percent of the total inequality is attributed to within-village inequality, suggesting that differences across villages can only explain about 40 percent of the total inequality. The decomposition over county village sample yields the similar results, as shown in the last row. Moreover, within-county and within-village inequalities both account for a larger proportion, which is over 10 percentage higher than in the full sample.

Our findings are consistent with several recent studies on spatial inequality in China. For example, using household data covering 18 provinces, Gustafsson and Li (2002) show that the contribution of inequality within counties (their smallest unit of analysis) to the total inequality is 57 percent in 1988 and 53 percent in 1995. Benjamin et al. (2002) find that only 25 to 30 percent of the rural inequality in 1995 comes from between-village differences, whichever inequality index they choose. However, their data are relatively small, with only 30 villages in Hebei and Liaoning provinces. Employing a data set from 9 provinces over 1986-1999, Benjamin et al. (2003) also conclude that more than half of the total inequality is driven by income differences within villages, and the proportion increases in recent years.

In Table 3 we further decompose each provincial inequality to explore the regional differences in income distribution. We find that within-village inequality contributes at least about one half to provincial inequality, although the contribution varies across provinces. Jiangsu province has the highest per capita income (1,560) but the second lowest proportion of within-village inequality (49 percent). In contrast, Henan province has the lowest per capita income (689) but the second largest contribution of within-village inequality (76 percent). For other provinces, Shandong has the smallest contribution of within-village inequality while Guangxi has the largest, although the income and inequality levels of the two provinces are very close. This suggests that some region-specific factors may play a role in the evolution of local inequality.

2.4 The Determinants of Income Inequality

In previous parts we show that over half of the rural inequality in China is attributed to income inequality within villages. In this subsection we focus our analysis on the village units to explore the potential determinants of village inequality, i.e., the distribution of household endowments, and village-level institutions and market development. We use OLS regressions to examine the effects of these factors on village inequality.

Following Benjamin et al. (2002), we hypothesize that village inequality is positively correlated with the inequality of household endowments such as human capital (education) and land. Intuitively, human capital plays an important role in non-farm earnings, as more skilled (educated) individuals are more likely to gain benefits from market development by working in TVEs, setting up their own business or migrating to urban areas. On the other hand, land is one of the key factors of rural production and thus significantly determines the agricultural income. However, since land is allocated based on household size under HRS, the distribution of per capita land may be more equal than other endowments.

Our second hypothesis is that village inequality may increase with better development of non-agricultural market and village institutions. Given the same distribution of human capital, more access to non-agricultural sector enables more individuals to engage in non-farm business. As non-farm income is a more disequalizing source than agricultural income, we may expect total inequality to rise. Furthermore, market development may also interact with the change of human capital in affecting village inequality.

Next, we test above hypotheses using OLS models, with village inequality as the dependent variable. Table 4 shows the descriptive statistics of the variables used in our regressions. The income, education and land variables are all derived from household data. The village Gini coefficient ranges from 0.13 to 0.56, with a mean of 0.34. Per capita education is calculated over the average formal schooling years of adult labors in households. The land variable is collected from household survey of farming income, but it may understate the

real land households own.[6] On average, the Gini of land is marginally higher than the Gini of education (0.24 versus 0.22).[7] Finally, the institution and market development variables are directly obtained from the separate community survey, except for the average share of wage income, which is also calculated based on household data.

Table 5 reports our estimation results for education and land. In addition to the Gini of each endowments, we also control for their levels, the average village income and area indicators. We find that the distribution of education has no significant effect on inequality (column 1). But the distribution of farming land is positively correlated with inequality, whether we use the non-missing sample (column 2) or the full sample (column 3). One-standard-deviation increase (0.67) in the Gini of land will increase the village inequality by 0.1. Even after we control for education, the coefficient on Gini of land is still significant (column 4). In all the specifications, the coefficients on per capita income are significantly negative, suggesting that inequality decreases with higher average income.

In Table 6, we test our second hypothesis regarding the effects of market development. Among the four variables we use, only the share of work force in TVEs is statistically significant but has the opposite sign pattern as we expect (column 4). The coefficients on other three variables have the expected sign, although none is significant (column 1 to 3). Similar to Table 5, the Gini of land and average income are significant through all the specifications. We further examine the effects of interaction between the Gini of education and market development variables in Table 7, but none of the interaction term turns out to be significant. The results are not satisfying probably because the development of non-agricultural sector and market institution is a complex process and may not be perfectly captured by the variables we choose. However, these may be the best variables available in our sample since the design of CHNS survey focuses more on nutrition and health parts.

[6]The survey question is "In last year, how many mu of land did your household cultivate?"

[7]Since in some villages the majority of households may specialize in agricultural activities other than farming, about 25 percent of our sample villages have missing value in land. To avoid the shrink of sample, we replace the missing observation with zero and generate a dummy indicating such a missing. The means of the full sample are therefore smaller.

3 Income, Income Inequality, and Health in China

In this section, we test several hypotheses regarding the relationship between income, income inequality and health, employing the CHNS data. The structure of the section is as follows. Subsection 3.1 presents the hypotheses and literature review. Subsection 3.2 describes the data and some measurement issues. Subsection 3.3 reports our estimation results.

3.1 Income, Income Inequality and Health: Hypotheses and Previous Literature

In our study we attempt to examine whether health outcomes and behaviors are correlated with income and income inequality in China. We begin with a discussion of several hypotheses that link income and income distribution to health, followed by a selected review of previous empirical work. We then specify the empirical test for each hypothesis.

Hypothesis 1: Absolute Income Hypothesis

The *absolute income hypothesis* argues that people with higher incomes have better health outcomes, but income inequality or relative income has no direct effect on health. A related concept is the *poverty hypothesis*, which emphasizes that ill health is a consequence of low income or extreme poverty. The idea that health improves with income goes back a long way in the literature. One of the most influential works in this area is by Preston (1975), who finds that the impact of additional income on mortality is greater among the poor than the rich. In other words, there is a concave relationship between income and health.

A large number of empirical studies in a variety of disciplines (such as economics, sociology and epidemiology) demonstrate a robust association between income and health (no matter how income and health are measured) using individual data, and most of the evidence points to a nonlinear relationship.[8] We follow the literature and test whether per

[8]See the review in Feinstein (1993), and a more recent discussion in Smith (1999).

capita income has a positive effect on individual health.[9] However, since the protective effect of absolute income on health is relatively uncontested (compared with the effect of income inequality or relative income), we do not place very much emphasis on this test.

Hypothesis 2: Income Inequality Hypothesis

The *income inequality hypothesis* presumes that income inequality *per se* is a threat to the health of individuals within a society, even holding their incomes constant. It focuses on the direct tie between health and income inequality, regardless of a person's particular income level. There are several potential pathways through which income inequality might harm an individual's health directly. For example, high levels of inequality might produce instabilities in the social capital, by, for example, increasing mistrust and stress, or declining social cohesion, which in turn adversely influence an individual's own health through psychosocial responses like violent crime or self-destructive behaviors.[10]

This hypothesis has two versions (Mellor and Milyo, 2002). The strong version states that inequality affects all members in a society equivalently, irrespective of their income levels. The weak version suggests that income inequality may harm the health of only the least well off in a society, or that the harmful effect of inequality on health decreases with one's income rank.

Early studies use aggregate data to test the correlation between income inequality and health. Various works by Wilkinson over the past decade (e.g., 1992, 1996) present evidence of a relationship between income inequality and life expectancy across a number of industrialized countries, both at a point in time and over time. While Wilkinson reports correlation coefficients, a growing body of literature tests this hypothesis using regression frameworks. A link between income inequality and health measures (mortality, morbidity, etc.) has been discerned repeatedly at the level of countries (Waldmann, 1992; Wennemo, 1993), and across states, counties and cities within nations (Kaplan et al., 1996; Ben-Shlomo

[9]We also control for income squared to capture the nonlinear relationship between income and health.

[10]Kawachi and Kennedy (1999) summarize three plausible mechanisms linking income inequality to health: disinvestment in human capital, the erosion of social capital, and stressful social comparisons.

et al., 1996; Kennedy et al., 1996; Kawachi et al., 1997; Kawachi and Kennedy, 1997; Lynch et al., 1998). In addition, some studies find an association between income distribution across U.S. states and state-level measures of smoking (Kaplan et al., 1996), alcohol consumption (Marmot, 1997) and firearm crimes (Kawachi et al., 1998).

Although these studies are informative, they use aggregate data, making it hard to differentiate between the hypotheses for absolute income and income inequality. The aggregate association between income inequality and health may merely reflect the nonlinear relationship between income and health at the individual level. For example, if a transfer of one dollar from the rich to the poor improves the health of the poor more than it diminishes the health of the rich, this income-equalizing transfer will increase the average health of the whole society.[11] If all that matters to individual health is income, then for two communities with identical average income, the community with a more equal income distribution tends to have better average health than the one with greater inequality. Thus, in aggregate studies, it is hard to distinguish this "statistical artefact" (Gravelle, 1998) from mechanisms in which income inequality has a direct effect on individual health. In order to identify the true effect of inequality, one should employ individual data.

A number of studies using U.S data find that income inequality does indeed have a negative effect on individual health. For instance, Kennedy et al. (1998), Soobader and LeClere (1999), Fiscella and Franks (2000), and Blakely et al. (2001) all show a significant association between inequality (at state or county-level) and self-rated health status. Daly et al. (1998) examine the effects of several measures of state-level income inequality on individual mortality, and find supporting evidence for the income inequality hypothesis in a particular time period. Using county and tract-level inequality data, LeClere and Soobade (2000) find supporting evidence as well, but only for some specific subgroups in high-inequality counties.

In contrast, some studies indicate no association between income inequality and individual health. Measuring inequality by the proportion of income earned by the poorest 50 percent of the population, Fiscella and Franks (1997) find no effects of county-level inequal-

[11] Using a new data set, Deaton (2003) shows a recent version of the Preston curve and suggests that income redistribution from rich to poor countries will in principle increase average health worldwide.

ity on mortality. Meara (1999) examines the relationship between state-level inequality and birth outcomes (such as infant mortality and low birth weight), and finds no significant relation. Mellor and Milyo (2002) construct several inequality measures both at the level of states and metropolitan areas, and show that their effects on self-rated health status are eliminated once individual income and locality effects are controlled. Using the same data as Mellor and Milyo (2002), Blakely et al. (2002) draw a similar conclusion, finding that, after controlling for income, there is little association between income inequality and individual health. A few studies using data outside the U.S. provide further evidence against the income inequality hypothesis (Osler et al., 2002; Shibuya et al., 2002; Gerdtham and Johannesson, 2004).

Most of the existing literature focuses on the strong version of the income inequality hypothesis. Only a few studies (Daly et al., 1998; Meara, 1999; Mellor and Milyo, 2002; Gerdtham and Johannesson, 2004) implicitly or explicitly test the weak version, but none of their findings support the hypothesis.

In this thesis, we test both the strong and weak versions of the inequality hypothesis. The strong version of the income inequality hypothesis is specified as follows,

$$H_{ij} = \beta_0 + \beta_1 Q_j + \beta_2 Q_j^2 + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}, \tag{3}$$

where i and j are subscripts for individual and community respectively. H_{ij} denotes a number of health outcomes and behaviors (self-reported health status, objective body conditions, smoking, alcohol use, etc.). Q_j stands for the community-level income inequality. I_{ij} is the vector of per capita income and income squared, and X_{ij} is the vector of other individual, household and community variables. We also include the squared term of inequality to capture the potential nonlinear effect. We hypothesize that health outcomes deteriorate with income inequality ($\beta_1 < 0$), but the relation might not be linear ($\beta_2 \neq 0$).

To test the weak version, we extend equation (3) by introducing the interaction between inequality and a person's rank (in ascending order of income), denoted by R_{ij}, to allow the effects of income inequality to vary by the relative income level. The model is

$$H_{ij} = \beta_0 + \beta_1 Q_j + \beta_2 Q_j^2 + \delta R_{ij} + \eta Q_j \cdot R_{ij} + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}. \tag{4}$$

We expect a positive coefficient on the interaction term ($\eta > 0$), or that the negative effect of inequality on health outcomes is smaller for people with higher income rankings.

Hypothesis 3: Relative Income Hypothesis

The *relative income hypothesis* states that health depends on an individual's income relative to others in his or her group, rather than an individual's absolute income. According to this hypothesis, health declines when one is financially deprived relative to one's peers, and improves when one is prosperous relative to others. A similar hypothesis is the *relative position hypothesis*, which stresses that one's relative rank in a group is related to health outcomes.[12]

Some psychosocial and material factors may play a role in the mechanisms connecting relative income to health. Perceptions of being relatively deprived compared to their peers may make people stressed and depressed, thus diminishing their health directly through diseases or indirectly via health-compromising behaviors.[13] Another possibility is that within a community, relative income (or rank) may be more important in determining an individual's access to material goods or services that are correlated with health.[14]

The relative income hypothesis is consistent with an effect of income inequality, but the two are not totally equivalent. If inequality increases, the poor are made even poorer in relative terms, and the rich become relatively more prosperous. Thus the harmful effect of income inequality is greater among the least well off. In this sense, the relative income theory parallels the weak version of the income inequality hypothesis. However, the strong version of the income inequality hypothesis goes further than the relative income hypothesis. According to the strong version, even rich people, who are least deprived in terms of relative income, may still suffer the adverse impacts of high income inequality. Thus, the strong

[12]The rank extends the concept of relative income as it can be measured by socioeconomic factors other than income, such as occupation and education.

[13]Some research on monkeys and primates (e.g., Cohen et al., 1997; Shively et al., 1997) provides biological evidence of how relative status may affect health.

[14]Deaton (2003) takes the case of local housing in a town: the richest people are able to get the hilltop plots with fine views while the poorest are left with the plots downward of the smokestacks. This is an example "where it is not money itself that is important, but *rank*, here determined by money."

version suggests that income inequality might directly influence health through channels independent of relative income.

Studies using different measures of relative income generate mixed results. Some recent research uses the mean (or median) income of a community as a proxy for relative income, but finds no evidence supporting the hypothesis (eg., Robert, 1998; Gerdtham and Johannesson, 2004). However, the Whitehall study in Britain (Marmot et al., 1984; Marmot et al., 1991), one of the most widely-known studies on relative income (position), finds higher rates of morbidity and mortality among civil servants in the lower administrative ranks. The contributions by Deaton (2001) and Eibner and Evans (2001) are more interesting, since they measure the level of relative income more specifically by the differences between an individual's income and the incomes of the richer members of the group. Using these measures, which are called *relative deprivation* (RD),[15] they both find a significant relative-income effect on individual mortality from U.S. data. Moreover, Eibner and Evans (2001) show that relative deprivation also influences the probability that an individual will engage in health-compromising behaviors, such as smoking and not wearing a seatbelt while driving.

Following Eibner and Evans (2001), we test the relative income hypothesis using the following specification,

$$H_{ij} = \beta_0 + \beta_1 RD_{ij} + I_{ij}\Gamma + X_{ij}\Theta + \varepsilon_{ij}. \tag{5}$$

Equation (5) is similar to equation (3), except that we replace Q_j with RD_{ij}, which stands for relative deprivation indices that measure an individual's relative income (see section 3.2.2 for details). The difference in subscripts between Q_j and RD_{ij} means that income inequality is an aggregate measure for the whole community, while the relative income measures that we use are individual-specific. We hypothesize that higher relative deprivation of income (or lower relative rank) reduces the probability of being healthy, and increases the probability of participating in health-compromising behaviors.

[15]The definition of relative deprivation is originally proposed by Runciman (1966), who argues that one is deprived if others in the group possess something that one does not have. Yitzhaki (1979) develops the definition by viewing income as personal possessions, and shows the link between relative deprivation and income inequality.

3.2 Data and Measurement Issues

In this thesis, we use the China Health and Nutrition Survey (CHNS) data, which were collected by the Carolina Population Center (CPC) at the University of North Carolina at Chapel Hill, the Institute of Nutrition and Food Hygiene, and the Chinese Academy of Preventive Medicine in 1993. The sample households were randomly drawn in eight provinces including Liaoning, Shandong, Jiangsu, Henan, Hubei, Hunan, Guangxi, and Guizhou. Two cities and four counties were sampled in each province. Four neighborhoods in each city, and one county-town neighborhood and three villages in each county, were then randomly selected. We define a neighborhood or village as a community unit. Approximately 20 households were sampled per community.

The CHNS data contain detailed information on household and individual characteristics as well as health-related information such as self-reported health status, physical conditions and health behaviors. We restrict our sample to men and women over the age of 20 for whom we have a complete set of data on health and demographic variables (age, sex, marital status, education, family income, etc.). As we need to construct income inequality and relative deprivation indices, we also exclude those with non-positive household income. In total, we have 7,300 observations in the sample.

Table 8 summarizes the definitions of variables in our sample. We now discuss a variety of measurement issues that need to be clarified before we present our estimation results.

3.2.1 Health Indicators

The CHNS data offer several potential health measures, as shown in the top panel of Table 8. Self-reported health status (SRHS) is the main health measure we use. Although SRHS is a subjective measure of individual health,[16] previous studies show that SRHS is highly correlated with subsequent mortality, even when controlling for more objective health measures (Idler and Benyamini, 1997; Deaton and Paxson, 1998). We construct a binary

[16]In the survey the interviewees were asked the question "right now, how would you describe your health compared to that of other people of your age."

variable, SRHS, which equals one if excellent or good health is reported and equals zero if fair or poor is reported.[17]

We also use several objective health measures such as physical functions (PF) and activities of daily living (ADL), which are recorded in the physical examination section of the survey. PF provides information on the status of various body functions associated with heart, hearing, eyesight, arms, legs, etc. We use two common indicators as PF measures: one is of heart, lungs and stomach condition, and the other of blood pressure. As with SRHS, we define two binary variables that equal one if the function is normal and zero otherwise. ADL measures whether or not the individual is physically restricted or unable to perform daily activities, such as taking a bath, eating and drinking alone, or putting on clothes. We create two binary variables that are equal to one if respondents are able to perform the walking and lifting activities respectively, and equal to zero if respondents report any difficulty in these activities. However, ADL measures are unavailable for individuals under fifty, thus we can only use this measure for a sample of 1,998 observations of the elderly.

Besides these direct measures, the CHNS data contain information on some health-compromising behaviors such as smoking and alcohol consumption. Regarding smoking behavior, we have knowledge of whether or not an individual smoked at the time of the survey, and the number of cigarettes smoked per day. Regarding drinking behavior, we know whether or not an individual had drunk any alcoholic beverage in the year prior to the survey, and the frequency of drinking. In total, we have four variables to measure health behaviors, i.e., current smoker, cigarettes per day, current drinker and drinking frequency, as illustrated in Table 8. As most of the smokers and drinkers are men in our sample, we generate a sub-sample of 3,172 observations, by limiting our sample to men who have non-missing behavior variables.

Table 9 provides descriptive statistics concerning these health measures. SRHS and PF measures are available for the whole sample, but ADL and health behavior variables are only available for smaller samples. Among all individuals, 73 percent reported being in good health. Examining the data in two sex groups, we find that men are more healthy

[17]In the survey, SRHS is a categorical variable coded on a scale of one (excellent) to four (poor).

than women, with 76 percent of men but only 70 percent of women reporting themselves in good health. The proportion declines with age, as only 56 percent of those over fifty report themselves to be in good health. By contrast, higher normal rates are reported for the two measures of physical functions, both exceeding 90 percent for the whole sample. The proportion of people with no limitations in daily activities is close to that for SRHS, although it should be remembered that the sample is much smaller. [18]Finally, 65 percent of men were smoking at the time of the survey, and 63 percent reported that they drank during the year prior to the survey.

3.2.2 Income Inequality and Relative Income Measures

In this thesis we mainly use the Gini Coefficient to measure the community-level income inequality.[19] For every community, we calculate the Gini based on household income weighted by the family size. In total there are about 180 communities in our sample. The Gini ranges from 0.1 to 0.6, with the average value around 0.32 (Table 9).

Following Eibner and Evans (2001), we construct several relative deprivation indices as the proxy for relative income, i.e., relative deprivation of absolute income (RDA), relative deprivation of log income (RDL), relative deprivation over individual income (RDI), and individual rank. Based on the theory developed by Yitzhaki (1979), RDA is defined as:

$$RDA_i = \frac{1}{N}\sum_j (y_j - y_i) \qquad \forall \quad y_j > y_i. \tag{6}$$

It measures the relative deprivation of person i with income y_i in a reference group of N people by the normalized total incomes of other group members who earn more than i does. RDL is the same as RDA except that it uses $log(y)$ rather than y in (6). RDI equals RDA_i/y_i, namely the ratio of RDA relative to person i's own income. The final index we use is the individual's centile rank within the reference group (where income is sorted in ascending

[18]The subjective SRHS measure and objective PF or ADL measures are positively correlated, as shown by their correlation coefficients in Table 10.

[19]Kawachi and Kennedy (1997) show that the six inequality measures (including the Gini Coefficient and the Theil index) used in their study are highly correlated with each other, and the choice of inequality indicators does not change the relationship between income inequality and mortality. We also use another inequality index, namely the Theil index, to test the robustness of our results, and find that using different measures of inequality does not change our results qualitatively.

order). In contrast to the first three measures, the rank ignores the magnitude of the income difference between individuals. While larger values in RDA, RDL and RDI indicate higher levels of relative deprivation, higher centile rank means a lower level.

As the Gini Coefficient depicts the overall income distribution of a society, relative deprivation reflects a person's position or rank relative to the incomes of others within a reference group. In order to be consistent with the Gini Coefficient, we use households in the same community as the reference group to generate these RD measures. The summary statistics of our relative deprivation measures are reported in Table 9. Unlike the Gini, which is bounded between 0 and 1, relative deprivation measures (RDA, RDL and RDI) are not limited in value and therefore have larger variations in the sample.

3.2.3 Other Explanatory Variables

In the individual-level analysis, we control for variables including per capita income and income squared, age and age squared, education, indicators for sex and marital status, family size, household environment, the distance from the community to nearby medical facilities, and year, rural and provincial indicators. We show the descriptive statistics for these variables in Table 9. Individuals in our sample have an average income of 1,374 yuan.[20] Household environment measures the degree of excreta contamination around the respondent's dwelling place and is directly recorded through the interviewer's own observation. The distance to medical facilities is obtained from the CHNS community survey and measures the availability of public health services to the community. We use the average distance if more than one facility is frequently used.

In Table 11 we divide the sample into two sub-samples: good health and poor health (columns 1 and 2). The differences in personal characteristics between the two sub-samples are what we would intuitively expect. Specifically, we find that on average healthy people have higher per capita income and education level, and are much younger than unhealthy ones. Those in good health also live in larger families, in better environments, and closer to

[20]We use the consumer price index included in the CHNS data to adjust per capita income to prices in urban areas in Liaoning province.

medical facilities. The role of income inequality is less explicit, as the average Gini Coefficient and Theil index for the two groups are very close. On the other hand, the poor health group on average is slightly more deprived, as indicated by its smaller mean of individual rank and larger mean of the other three indices. The t-ratios in column 3 show that most of the means are significantly different between the two sub-samples, except for some inequality and relative deprivation variables.

3.3 Estimation Results

In this subsection we employ OLS and Probit models to systematically test various hypotheses discussed in Subsection 3.1. The main purpose of our study is to examine the correlation between individual health and income inequality or relative income. We also make changes in the model specifications to test the robustness of our results.

3.3.1 Income, Income Inequality and Individual Health

We first employ Probit and OLS models to test the income inequality hypothesis (Hypothesis 2), in both the strong and weak versions. We apply models (3) and (4) to various health measures such as SRHS, PF, ADL and health behaviors, using individual-level data. We use the Gini Coefficient as the inequality index in this subsection.[21] Our specifications also allow us to test the absolute income hypothesis (Hypothesis 1), even though it is not our focus.

Self-Reported Health Status

Table 12 presents the results of probit regressions using SRHS as the dependent variable. The results exhibit an "inverted-U," i.e., a quadratic relationship between SRHS and income inequality. We report dF/dx, or the marginal change of probability of reporting excellent or

[21] As a robustness check, we repeat all the regressions using the Theil index and obtain very similar results.

good health when the independent variable increases. In the first column, we have the Gini as the only independent variable. The coefficient on the Gini is positive but not significant. When we add the squared term in the second column, the correlation is still insignificant. However, in column 3, the coefficients on the Gini and Gini squared both become significant at the five percent level, after we include other control variables such as per capita income, and personal and household characteristics. The positive coefficient on the Gini and negative coefficient on Gini squared mean that SRHS increases with inequality when Gini is less than 0.40 (75 percentile in the sample) and decreases with inequality for larger Gini. The results suggest that the strong version of the income inequality hypothesis (Hypothesis 2) is only supported for communities with large inequality.[22]

We also find evidence supporting the absolute income hypothesis (Hypothesis 1). Column 3 shows that there is a concave relationship between individual health and per capita income. The positive coefficient on income and negative coefficient on income squared are both significant at the one percent level. The critical point of the health-income quadratic curve is about 6,833 yuan, but 99 percent of the values for income in our sample are below this figure. This means that for most of our sample health increases with absolute income, but at a decreasing rate.

Other control variables also have the expected signs in column 3. The probability of being in good health decreases with age at a rate of 1.1 percentage points per year. One more year of schooling increases the probability of being in good health by 0.3 percentage points. Men have a 3.9 percentage points higher probability of being in good health than women, and married people have a 4.2 percentage points higher probability than single people. A one-standard-deviation increase in family size (1.6) raises the probability by 1.9 percentage points. The protective effect of good household environment is sizable, increasing the probability of reporting good health by 13 percentage points. Finally, the distance to medical facilities has a negative sign but it is statistically insignificant.

Next in column 4, we test the weak version of the income inequality hypothesis, i.e., whether the effects of inequality differ by relative income. As in the previous regression, the

[22]This is consistent with the findings of LeClere and Soobade (2000) who use US data.

Gini has a quadratic effect on health. Moreover, the interaction between the Gini and the individual rank is negative and significant, which suggests that the partial effect of the Gini depends on both the rank and the Gini itself. For example, at the mean level of the Gini (0.32) and the individual rank (0.50), the total partial effect of income inequality on health is $1.051 + (-1.065 \times 2 \times 0.32) + (-0.378 \times 0.50) = 0.180$. This means that an increase in the Gini by one standard deviation (0.10) will lead to a 1.8 percentage points higher probability of reporting good health. This effect decreases with inequality and becomes negative at high levels of inequality. But the negative interaction suggests that, for people with higher rankings, the detrimental effect of income inequality on health is greater. This result seems to contradict what is predicted by the weak version of the income inequality hypothesis, that income inequality harms the health of the poor more than the rich.

However, the result is more interesting if we examine the partial effects of the individual rank. The coefficient on the rank is significantly positive. Thus we can interpret the negative interaction as implying that living in a more unequal community would dilute the positive effect of the rank on health. At the mean level of the Gini (0.32), the total partial effect of an increase in the rank on health is $0.148 + (-0.378 \times 0.32) = 0.027$, but the effect decreases with the Gini. The enhancing effect of personal rank on health becomes smaller with more inequality, and even turns negative when inequality is very high (the Gini above 0.39).

In short, the results in Table 12 show that the community-level income inequality influences the individual health status in a nonlinear way. According to the estimated coefficients, income inequality tends to have a detrimental impact on health when a community has large inequality (the Gini above 0.40, in column 3). The higher individual rank is beneficial to one's health, but this positive effect becomes weaker in a more unequal community.

Physical Functions

Table 13 reports estimations using two PF variables as dependent variables: the condition of heart, lungs and stomach, and the condition of blood pressure. We find a nonlinear

relationship between the Gini and heart function (columns 1-3), but no correlation between the Gini and blood pressure (columns 4 and 5). The effects are not altered by one's relative income position, as the coefficients on the interaction term are insignificant (columns 3 and 6). Only a few of the control variables are significant.[23]

Activities of Daily Living

As another check, we estimate the influence of income inequality on ADL measures in a restricted sub-sample of elderly people. The two dependent variables we use are indicators of whether one is able to walk for one kilometer and lift a five-kilogram bag without difficulty. We follow the estimation specifications that were previously applied to PF indicators.

The regression results in Table 14 further confirm our finding that income inequality has an impact on individual health. The community Gini has a negative effect on both walking and lifting abilities (columns 1 and 4). Moreover, inequality has a nonlinear effect on the lifting ability (columns 5 and 6). The estimation implies that the probability of being able to lift the bag decreases with income inequality when the Gini is greater than 0.29 (about 38 percentile in the sub-sample). The impacts of income inequality on ADL limitations are independent of the individual rank, since the interaction of Gini and rank is not significant in columns 3 and 6. Like PF variables, ADL measures may not be directly determined by general characteristics, because few of the control variables are significant in Table 14.

Health Behaviors

Previous results show that income inequality is strongly correlated to health outcomes. We now explore one of the potential mechanisms of their correlation by examining whether an increase in income inequality increases the probability that an individual engages in health-compromising behaviors, i.e., smoking and alcohol consumption. The probit and

[23]One concern about the above results is that PF measures may lack variation in our sample. The proportion of people reporting normal heart condition amounts to 93 percent, and the proportion reporting normal blood pressure is 95 percent.

OLS regression results using different dependent variables are reported in Table 15.

Table 15 (columns 1 and 2) shows a strong correlation between inequality and smoking habits. In the first column we have the current smoker indicator as the dependent variable. The coefficient on the Gini is positive and significant at the five percent level. It predicts that a one standard deviation increase in community Gini (0.10) will increase the probability of smoking by 2.6 percentage points. We then use the OLS model to estimate the effects on cigarettes consumed per day in the second column. As with the estimation on current smoker, the Gini has a strong positive effect.

Table 15 (columns 3 and 4) also exhibits a strong association between inequality and drinking behavior. The effect of income inequality on the probability of being a current drinker is positive and significant at the five percent level. The pattern is similar for drinking frequency. Specifically, the coefficients on the Gini suggest that a rise in the Gini by one standard deviation (0.10) causes a 2.2 percent higher probability of drinking alcohol, and an increase in drinking frequency by approximately half of its standard deviation (2.01).

3.3.2 Relative Income and Individual Health

We now test the relative income theory (Hypothesis 3) by replacing the independent variables of inequality with relative deprivation measures: RDA, RDL, RDI and individual rank. The model to be estimated is equation (5). Because these measures are highly correlated with each other, their effects are estimated separately.

The estimation results with SRHS as the dependent variable (Table 16) show that the relative income hypothesis is only supported when relative deprivation is measured by one's income rank. In columns 1 to 3, none of the coefficients on RDA, RDL and RDI is statistically significant. On the other hand, the individual rank has a significantly positive effect on SRHS, even after we control for absolute income (column 4).[24] Holding an individual's income constant, increases in other people' income (thus lowering the individual's own rank) can be harmful to the individual's health.

[24]This effect is already shown when we test the weak version of the income inequality hypothesis (Table 12, column 4).

We conduct the same estimations taking PF/ADL and health behavior measures as dependent variables, but do not find any significant correlations with the relative deprivation indices, including the individual rank (hence the results are not reported). Our results differ from those of Eibner and Evans (2001), who find that the relative deprivation has a stronger impact on health when it reflects income differences between individuals (measured in RDA, RDL and RDI). Their results are imprecise in many cases when they measure relative deprivation using rank. However, our results may be sensitive to the reference group we define.[25]

[25]Due to the relatively small sample (about 40 individuals per community on average), we are not allowed to define narrower reference groups by age or education within the community, as Eibner and Evans (2001) are able to do.

4 Conclusion

In this thesis we employ micro data from China to explore several aspects of income inequality in China. We first examine the patterns and determinants of income inequality. We find that over half of the total inequality is driven by within-region inequality. However, our cross-section analysis do not find consistent evidence that the two categories of factors outlined in Benjamin et al. (2003) (i.e., distribution of household endowments, and village-level market development and institutions) play important roles in the evolution of within-village inequality.

We also investigate one particular socioeconomic impact of income inequality, by testing several hypotheses linking income and income inequality to individual health status. We find some evidence supporting these hypotheses. First, our results show a concave relationship between self-reported health status and per capita income (the absolute income hypothesis). Additional income brings about greater improvement in the health of the poor than of the rich. Second, we find a significant association between self-reported health status and community-level income inequality (the income inequality hypothesis). In fact, the relationship we find appears as an "inverted-U" shape. That is to say, rising inequality tends to improve health when inequality is low, and to harm health when inequality is above a certain level. We also find evidence that income inequality may influence health via increasing the likelihood and frequency of health-compromising behaviors such as smoking and alcohol consumption. Finally, the centile rank of income has a strong positive effect on self-reported health status (the relative income hypothesis), but its protective effect decreases with inequality and turns negative under extremely high inequality.

While this study has its own limitations, it is among the first to provide evidence from a developing country on the negative association between inequality and health, both of which are important issues for students of development. Although the sample size is relatively small compared with the data in many U.S. studies, the set of CHNS data we have used is so far one of the best data sets used in studying inequality and health in the context of developing economies, and is probably the best Chinese data set. Another limitation is

that we only focus on one dimension of inequality, i.e., community-level inequality. We do not claim that community-level inequality is necessarily more important than inequality at county- or provincial-level; rather, our purpose is to examine the socioeconomic impacts of inequality in a local setting, where we can see the people interacting with each other more closely. Focusing on the community level can also facilitate the empirical tests by allowing a larger variation of inequality in the sample. Finally, strictly speaking, our empirical tests are tests of correlations between community-level inequality and individual health. The causal link may not be established until more evidence becomes available regarding the intermediate mechanisms through which inequality affects health. However, intuitively, the causality is more likely to go from inequality to health because it would be difficult to argue that individual health affects the community-level inequality.

China began its economic reform by abandoning the principle of absolute equality, "eating from the same kitchen system," in agriculture (Lin, 1992), in industry (Li, 1997) and even in government (Qian and Weingast, 1997). The reforms have improved incentives in most workplaces, which in turn has led to historic levels of growth in the past 25 years. However, the ever-increasing inequality that accompanies growth will ultimately slow it down. A recent study by Benjamin et al. (2004) finds that village-level inequality is negatively associated with village economic growth in the long run. While there are many channels through which inequality could affect growth, this thesis shows a particular one, poor health, which is itself a direct indicator of underdevelopment.

The Chinese government has apparently taken note of the serious issue of inequality. Wen Jiabao, the new premier, has repeatedly told the public that the goal of this government is to achieve equitable growth. The government has recently been shifting its focus from the more developed coastal areas to the poor inland areas, introducing a series of preferential policies in favor of the latter, such as a wider range of fiscal subsidies, lower tax rates and cheaper loans. The government is also shifting its focus from the fast developing industries to the sluggish agricultural sector which employs most of China's poor. It plans to remove all agricultural taxes in the next two to three years. While many policies are aiming at

decreasing the inter-region disparity, the government should also put equal wight on policies that may reduce local inequality at the same time. As suggested by our results, elimination of within-region inequality will contribute a large proportion to the decrease of total inequality, and income redistribution will improve the health of the population, especially in regions where large inequality prevails.

Table 1: Source and Distribution of Household Income in Rural China, 1993

	Mean (Overall)	Percentage Not Zero	Mean (Not Zero)	Gini Coefficient
Full Sample (N=2,276)				
Source of Income				
Agriculture				
Gardening	705.3	60.9	1159.0	0.60
Farming	565.4	47.5	1190.5	0.51
Livestock/Poultry	294.1	50.5	582.6	0.54
Fishing	15.5	2.1	752.8	0.66
Non-agriculture				
Wage Income	1112.7	38.9	2858.3	0.42
Small Household Business	713.6	19.0	3759.5	0.51
Welfare Subsidy	351.0	22.9	1530.3	0.51
Other Income	479.5	41.7	1148.9	0.61
Total Household Income	4758.2	100	4758.2	0.47
Per Capita Income	1187.2	100	1187.2	0.46
Sample Excluding Suburban Villages (N=1,755)				
Source of Income				
Agriculture				
Gardening	614.4	65.1	944.3	0.59
Farming	654.0	55.3	1183.3	0.50
Livestock/Poultry	302.0	57.1	529.0	0.52
Fishing	13.4	1.4	983.3	0.62
Non-agriculture				
Wage Income	819.9	32.0	2560.5	0.42
Small Household Business[1]	776.2	21.1	3671.9	0.51
Welfare Subsidy	145.0	14.6	990.4	0.50
Other Income[2]	330.8	37.4	885.1	0.60
Total Household Income	3995.4	100	3995.4	0.46
Per Capita Income	975.1	100	975.1	0.45

All incomes are in RMB yuan.

[1]Small household business includes handicraft or commercial business (such as carpentry, hairdressing, restaurant and store).

[2]Other income includes rental income, remittances, pensions and gifts, etc.

Table 2: Total Inequality and Its Decomposition in Rural China, 1993

	Total Inequality	Within-Province	Within-County	Within-Village
Full Sample				
(N=2,276)				
Gini Coefficinet	0.46			
90/10 Ratio	11.4			
Theil Index	0.368	0.343	0.249	0.213
(%)	(100)	(93.2)	(67.7)	(57.9)
Sample Excluding Suburban Villages				
(N=1,755)				
Gini Coefficinet	0.45			
90/10 Ratio	10.6			
Theil Index	0.351	0.326	0.283	0.245
(%)	(100)	(92.9)	(80.6)	(69.8)

Table 3: Provincial Inequality and Its Decomposition in Rural China, 1993

	Total	Within-County	Within-Village
Liaoning (N=275)			
Per Capita Income (yuan)	1342.2		
Theil Index	0.346	0.290	0.244
(%)	(100)	(83.8)	(70.5)
Jiangsu (N=276)			
Per Capita Income (yuan)	1559.8		
Theil Index	0.346	0.189	0.170
(%)	(100)	(54.6)	(49.1)
Shangdong (N=275)			
Per Capita Income (yuan)	1127.2		
Theil Index	0.331	0.210	0.152
(%)	(100)	(63.4)	(45.9)
Henan (N=244)			
Per Capita Income (yuan)	688.5		
Theil Index	0.362	0.302	0.276
(%)	(100)	(83.4)	(76.2)
Hubei (N=292)			
Per Capita Income (yuan)	1123.6		
Theil Index	0.297	0.208	0.156
(%)	(100)	(70.0)	(52.5)
Hunan (N=294)			
Per Capita Income (yuan)	1456.9		
Theil Index	0.429	0.303	0.267
(%)	(100)	(70.6)	(62.2)
Guangxi (N=303)			
Per Capita Income (yuan)	1204.0		
Theil Index	0.327	0.287	0.258
(%)	(100)	(87.8)	(78.9)
Guizhou (N=317)			
Per Capita Income (yuan)	957.0		
Theil Index	0.301	0.221	0.193
(%)	(100)	(73.4)	(64.1)

Table 4: Descriptive Statistics of Village Variables in Rural China, 1993 (N=122)

Variables	Mean	Standard Deviation	Min	Max
Income, Education and Land				
Per Capita Income (1000 yuan)	1.172	0.694	0.224	4.036
Gini of Income	0.344	0.093	0.132	0.559
Per Capita Education (years)	5.899	1.421	2.667	12.122
Gini of Education	0.222	0.070	0.082	0.432
Per Capita Land (non-missing, N=93)	1.132	0.662	0.235	4.632
Gini of Land (non-missing, N=93)	0.244	0.129	0.045	0.673
Land Missing Indicator	0.238	0.427	0	1
Per Capita Land (mu) [1]	0.863	0.753	0	4.632
Gini of Land[1]	0.186	0.153	0	0.673
Market Development and Institution				
Close to County Seat (km)[2]	-2.046	2.390	-13	0
Average Share of Wage Income	0.373	0.169	0	0.844
Share of Work Force out of Village[3]	0.210	0.214	0	0.95
Share of Work Force in TVEs	0.288	0.332	0	1
Suburban Indicator	0.238	0.427	0	1

[1]The mean for the full sample is calculated after replacing the missing values with zeros.

[2]We define the variable as the negative value of the distance from a village to the county seat. The distance for a suburban village is defined zero.

[3]The share of work force that worked out of village for more than a month during the year prior to the survey.

Table 5: OLS Regressions Measuring the Effects of Education and Land on Village Inequality

	Dependent Variable: Village Gini Coefficient			
	(1)	(2)	(3)	(4)
Gini of Education	-0.061			-0.136
	(-0.37)			(-0.82)
Per Capita Education	-0.013			-0.016*
	(-1.41)			(-1.75)
Gini of Land		0.139**	0.148**	0.154**
		(2.06)	(1.98)	(2.05)
Per Capita Land		-0.016	-0.017	-0.020
		(-1.22)	(-1.17)	(-1.35)
Per Capita Income (/1000)	-0.028*	-0.046**	-0.042***	-0.038**
	(-1.82)	(-2.15)	(-2.65)	(-2.29)
Land Missing Indicator			0.046	0.049
			(1.43)	(1.54)
Suburban Indicator	-0.025	0.041	-0.037	-0.033
	(-1.10)	(1.59)	(-1.57)	(-1.40)
Provincial indicators	Yes	Yes	Yes	Yes
Observation	122	93	122	122
R-squared	0.28	0.33	0.30	0.32

Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 6: OLS Regressions Measuring the Effects of Market Development on Village Inequality

	Dependent Variable: Village Gini Coefficient			
	(1)	(2)	(3)	(4)
Close to County Seat	0.001			
	(0.31)			
Average Share of Wage Income		0.017		
		(0.33)		
Share of Work Force out of Village			0.023	
			(0.61)	
Share of Work Force in TVEs				-0.088***
				(-2.79)
Gini of Education	0.022	-0.152	-0.096	-0.118
	(0.12)	(-0.92)	(-0.56)	(-0.72)
Per Capita Education	-0.006	-0.018**	-0.019*	-0.011
	(-0.53)	(-2.02)	(-1.97)	(-1.12)
Gini of Land	0.173**	0.143*	0.158**	0.166**
	(2.09)	(1.87)	(2.04)	(2.21)
Per Capita Land	-0.016	-0.020	-0.020	-0.031**
	(-1.01)	(-1.24)	(-1.30)	(-2.05)
Per Capita Income (/1000)	-0.035*	-0.038**	-0.037**	-0.030*
	(-1.92)	(-2.19)	(-2.12)	(-1.80)
Land Missing Indicator	0.054	0.042	0.054	0.057*
	(1.58)	(1.31)	(1.62)	(1.76)
Suburban Indicator	Yes	Yes	Yes	Yes
Provincial Indicators	Yes	Yes	Yes	Yes
Observation	112	119	118	115
R-squared	0.29	0.33	0.33	0.38

Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 7: OLS Regressions Measuring the Effects of Education and Market Development on Village Inequality

	Dependent Variable: Village Gini Coefficient			
	(1)	(2)	(3)	(4)
Close to County Seat	0.004			
	(0.35)			
Close to County Seat x	-0.014			
Gini of Education	(-0.27)			
Average Share of Wage Income		0.149		
		(0.81)		
Share of Wage Income x		-0.612		
Gini of Education		(-0.75)		
Share of Work Force out of Village			0.085	
			(0.65)	
Share out of village x			-0.307	
Gini of Education			(-0.50)	
Share of Work Force in TVEs				-0.177
				(-1.63)
Share in TVEs x				0.419
Gini of Education				(0.86)
Gini of Education	-0.011	0.061	-0.039	-0.159
	(-0.05)	(0.18)	(-0.19)	(-0.93)
Per Capita Education	-0.006	-0.019**	-0.019**	-0.008
	(-0.52)	(-2.10)	(-1.99)	(-0.79)
Gini of Land	0.172**	0.143*	0.166**	0.159**
	(2.07)	(1.87)	(2.09)	(2.10)
Per Capita Land	-0.016	-0.019	-0.021	-0.030**
	(-1.00)	(-1.23)	(-1.35)	(-1.99)
Per Capita Income (/1000)	-0.034*	-0.035**	-0.036**	-0.034*
	(-1.86)	(-2.04)	(-2.09)	(-1.96)

Land Missing Indicator	0.053	0.043	0.057*	0.061*
	(1.49)	(1.35)	(1.68)	(1.86)
Suburban Indicator	Yes	Yes	Yes	Yes
Provincial Indicators	Yes	Yes	Yes	Yes
Observation	112	119	118	115
R-squared	0.29	0.33	0.33	0.38

Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 8: Definitions of Health, Inequality and Other Variables in China

Variables	Definition
Self-Reported Health Status (SRHS)	1 if health is excellent or good, 0 if fair or poor
Physical Functions (PF)	
Heart Function	1 if normal in condition of heart, lungs and stomach, 0 if otherwise
Blood Pressure	1 if normal in blood pressure, 0 if with high blood pressure
Activities of Daily Living (ADL)	
Walking	1 if able to walk for a kilometer, 0 if with limitation
Lifting	1 if able to lift a 5-kilogram bag, 0 if with limitation
Health Behaviors	
Current Smoker	1 if smoke at the survey time, 0 if not
Cigarettes Per Day	Average number of cigarettes smoked per day
Current Drinker	1 if drink alcoholic beverage in the year prior to the survey, 0 if not
Drinking Frequency (0~5)	0 if not drink, 1 if no more than once a month, 2 if once or twice a month, 3 if once or twice a week, 4 if 3-4 times a week, 5 if daily or almost everyday
Inequality and Relative Deprivation	
Community Gini	Gini Coefficient of income within the community
Community Theil	Theil index of income within the community
Individual Rank	Centile rank (in ascending order of income) within the community
RDA	Yitzhaki's relative deprivation index: $RDA_i = (y_j - y_i)/N$, for all $y_j > y_i$, where y_i is income of person i and N is the size of the community
RDL	Substituting log(y) for y in RDA
RDI	RDA/y, i.e., dividing RDA by one's own income
Other Variables	
Income	Per capita household income
Education	Years of formal schooling
Age	(Restricting our sample to adults over the age of 20)
Male Indicator	1 if male
Married Indicator	1 if married, 0 if never married or divorced or widowed
Family Size	Number of household members, including adults and children
Household Environment	1 if little or no excreta around dwelling place, 0 if some or much
Distance to Medical Facility	Average distance of the community to most frequently used facilities
Rural Indicator	1 if the community is a village unit

Table 9: Descriptive Statistics of Health, Inequality and Other Variables in China

Variables	Mean	Standard Deviation	Min	Max
Self-Reported Health Status (N=7,300)				
Whole Sample	0.730	0.444	0	1
All Men	0.758	0.429	0	1
All Women	0.703	0.457	0	1
Age above 50	0.556	0.497	0	1
Physical Functions (N=7,300)				
Heart Function	0.928	0.259	0	1
Blood Pressure	0.947	0.224	0	1
Activities of Daily Living (N=1,998)				
Walking	0.671	0.470	0	1
Lifting	0.726	0.446	0	1
Health Behaviors (N=3,172)				
Current Smoker	0.650	0.477	0	1
Cigarettes Per Day	10.226	10.057	0	60
Current Drinker	0.634	0.482	0	1
Drinking Frequency	2.307	2.010	0	5
Inequality and Relative Deprivation (N=7,300)				
Community Gini	0.323	0.099	0.124	0.596
Community Theil	0.203	0.137	0.025	0.762
Individual Rank	0.498	0.303	0	1
RDA (/1000)	0.429	0.409	0	3.004
RDL	0.379	0.513	0	9.198
RDI	1.224	5.066	0	106.05
Other Variables (N=7,300)				
Income (1000 yuan)	1.374	1.247	0.0001	13.549
Education	6.050	4.381	0	18
Age	43.534	14.863	20	93
Male Indicator	0.498	0.500	0	1
Married Indicator	0.833	0.373	0	1
Family Size	4.415	1.590	1	13
Household Environment	0.804	0.397	0	1
Distance to Medical Facility (km)	1.494	2.764	0	22
Rural Indicator	0.677	0.468	0	1

Table 10: Correlations between SRHS, PF and ADL in China

Self-Reported Health Status (SRHS) and Physical Functions (PF) (N=7,300)

	SRHS	Heart Function	Blood Pressure
SRHS	1.000	-	-
Heart Function	0.158	1.000	-
Blood Pressure	0.126	0.142	1.000

Self-Reported Health Status (SRHS) and Activities of Daily Living (ADL) (N=1,998)

	SRHS	Walking	Lifting
SRHS	1.000	-	-
Walking	0.165	1.000	-
Lifting	0.146	0.475	1.000

Table 11: Descriptive Statistics of Variables for Healthy versus Unhealthy People in China (N=7,300)

	Mean and Standard Deviation		T-Statistics
	(1)	(2)	(3)
Variables	SRHS=1	SRHS=0	
Observations	5,332	1,968	
Inequality and Relative Deprivation			
Community Gini	0.323	0.322	0.57
	(0.098)	(0.100)	
Community Theil	0.203	0.202	0.37
	(0.136)	(0.139)	
Individual Rank	0.508	0.471	4.58***
	(0.304)	(0.301)	
RDA (/1000)	0.423	0.443	1.78*
	(0.415)	(0.392)	
RDL	0.368	0.407	2.90***
	(0.515)	(0.506)	
RDI	1.177	1.353	1.32
	(4.666)	(6.015)	
Other Variables			
Income (1000 yuan)	1.410	1.277	4.04***
	(1.270)	(1.178)	
Education	6.533	4.740	15.78***
	(4.220)	(4.540)	
Age	40.951	50.531	25.50***
	(13.701)	(15.619)	
Male Indicator	0.516	0.448	5.21***
	(0.500)	(0.497)	
Married Indicator	0.845	0.802	4.29***
	(0.362)	(0.398)	
Family Size	4.464	4.281	4.37***
	(1.542)	(1.707)	
Household Environment	0.832	0.729	9.91***
	(0.374)	(0.445)	
Distance to Medical Facility (km)	1.437	1.649	2.89***
	(2.590)	(3.185)	
Rural Indicator	0.686	0.650	2.91***
	(0.464)	(0.477)	

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 12: Probit Regressions Measuring the Effects of Income Inequality on Self-Reported Health Status

	Dependent Variable: Self-Reported Health Status (1=excellent or good, 0=fair or poor)			
	(1)	(2)	(3)	(4)
Community Gini	0.030	0.309	0.926***	1.051***
	(0.57)	(1.07)	(2.87)	(3.15)
Gini Squared		-0.414	-1.131**	-1.065**
		(-0.98)	(-2.47)	(-2.32)
Individual Rank				0.148**
				(2.41)
Gini * Rank				-0.378**
				(-2.10)
Control Variables				
Income (/1000)			0.041***	0.033***
			(4.40)	(2.66)
Income Squared			-0.003***	-0.002*
			(-2.70)	(-1.87)
Education			0.003**	0.003**
			(2.00)	(1.98)
Age			-0.011***	-0.011***
			(-4.51)	(-4.51)
Age Squared (/1000)			0.032	0.032
			(1.33)	(1.34)
Male Indicator			0.039***	0.039***
			(3.53)	(3.52)
Married Indicator			0.042**	0.043***
			(2.54)	(2.58)
Family Size			0.012***	0.012***
			(3.48)	(3.57)

Household Environment			0.129***	0.130***
			(8.89)	(8.92)
Distance to Medical Facility			-0.001	-0.001
			(-0.61)	(-0.69)
Rural Indicator			0.029**	0.027**
			(2.19)	(2.05)
Provincial Indicators	No	No	Yes	Yes
Observation	7300	7300	7300	7300
Pseudo R-squared	0.00	0.00	0.10	0.10

Note: Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 13: Probit Regressions Measuring the Effects of Income Inequality on Physical Functions

	Dependent Variable: Heart Function (1=Normal in heart, lungs and stomach, 0=Otherwise)			Dependent Variable: Blood Pressure (1=Normal blood pressure, 0=High blood pressure)		
	(1)	(2)	(3)	(4)	(5)	(6)
Community Gini	0.036	0.402**	0.430**	0.012	-0.076	-0.068
	(1.08)	(2.37)	(2.43)	(0.73)	(-0.92)	(-0.79)
Gini Squared		-0.533**	-0.542**		1.133	0.128
		(-2.20)	(-2.23)		(1.08)	(1.04)
Individual Rank			0.003			-0.003
			(0.08)			(-0.18)
Gini * Rank			-0.033			-0.005
			(-0.36)			(-0.10)
Control Variables						
Income (/1000)	0.005	0.005	0.008	-0.006***	-0.006***	-0.005
	(0.99)	(1.06)	(1.19)	(-2.61)	(-2.61)	(-1.55)
Income Squared	-0.0004	-0.0004	-0.001	0.0003	0.0003	0.0002
	(-0.70)	(-0.71)	(-0.90)	(1.16)	(1.12)	(0.68)
Education	0.0001	0.0003	0.0002	-0.0004	-0.0004	-0.0004
	(0.15)	(0.31)	(0.27)	(-0.97)	(-1.06)	(-1.10)
Age	-0.007***	-0.007***	-0.007***	-0.006***	-0.006***	-0.006***
	(-5.01)	(-4.95)	(-4.95)	(-6.91)	(-6.93)	(-6.93)
Age Squared (/1000)	0.042***	0.041***	0.041***	0.036***	0.036***	0.036***
	(3.19)	(3.13)	(3.13)	(4.82)	(4.83)	(4.84)
Male Indicator	0.013**	0.012**	0.012**	0.003	0.003	0.003
	(2.11)	(2.03)	(2.04)	(1.01)	(1.02)	(1.05)
Married Indicator	0.007	0.007	0.007	-0.008*	-0.007*	-0.007*
	(0.80)	(0.73)	(0.73)	(-1.90)	(-1.88)	(-1.88)
Family Size	0.0002	0.0001	0.00002	0.0001	0.0002	0.0002
	(0.12)	(0.03)	(0.01)	(0.14)	(0.20)	(0.18)

Household Environment	0.015*	0.015*	0.015*	0.006	0.006	0.006
	(1.79)	(1.80)	(1.78)	(1.39)	(1.39)	(1.34)
Distance to Medical Facility	-0.002**	-0.002**	-0.002**	0.0003	0.0003	0.0004
	(-2.20)	(-2.30)	(-2.24)	(0.48)	(0.56)	(0.58)
Rural Indicator	0.015**	0.012	0.012*	0.015***	0.016***	0.016***
	(2.02)	(1.59)	(1.65)	(4.17)	(4.26)	(4.30)
Provincial Indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observation	6359	6359	6359	6048	6048	6048
Pseudo R-squared	0.08	0.08	0.08	0.21	0.21	0.21

Note: Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 14: Probit Regressions Measuring the Effects of Income Inequality on Activities of Daily Living

	Dependent Variable: Walking (1=Able to walk for 1 km, 0=Have limitation)			Dependent Variable: Lifting (1= Able to lift a 5-kg bag, 0=Have limitation)		
	(1)	(2)	(3)	(4)	(5)	(6)
Community Gini	-0.549***	0.192	0.313	-0.333***	1.185*	1.246*
	(-3.73)	(0.24)	(0.37)	(-2.85)	(1.88)	(1.89)
Gini Squared		-1.062	-0.991		-2.174**	-2.143**
		(-0.93)	(-0.86)		(-2.45)	(-2.41)
Individual Rank			0.235			0.101
			(1.59)			(0.88)
Gini * Rank			-0.438			-0.204
			(-1.03)			(-0.61)
Control Variables						
Income (/1000)	0.003	0.004	-0.026	0.028	0.030*	0.019
	(0.15)	(0.17)	(-0.88)	(1.59)	(1.71)	(0.78)
Income Squared	0.001	0.001	0.003	-0.001	-0.001	0.0004
	(0.42)	(0.43)	(1.15)	(-0.25)	(-0.24)	(0.14)
Education	0.003	0.003	0.003	-0.003	-0.002	-0.002
	(0.65)	(0.72)	(0.71)	(-0.81)	(-0.60)	(-0.63)
Age	0.024	0.025	0.025	-0.011	-0.010	-0.009
	(1.19)	(1.21)	(1.25)	(-0.65)	(-0.58)	(-0.55)
Age Squared (/1000)	-0.311**	-0.314**	-0.320**	-0.072	-0.081	-0.083
	(-2.06)	(-2.07)	(-2.11)	(-0.57)	(-0.64)	(-0.65)
Male Indicator	0.126***	0.125***	0.124***	0.153***	0.152***	0.152***
	(4.44)	(4.40)	(4.36)	(6.76)	(6.70)	(6.71)
Married Indicator	-0.007	-0.008	-0.009	0.035	0.034	0.034
	(-0.21)	(-0.25)	(-0.27)	(1.29)	(1.24)	(1.25)
Family Size	0.012*	0.012*	0.013*	0.005	0.004	0.004
	(1.71)	(1.68)	(1.81)	(0.83)	(0.74)	(0.78)

Household Environment	0.045	0.043	0.047	0.043	0.040	0.041
	(1.27)	(1.23)	(1.32)	(1.49)	(1.39)	(1.43)
Distance to Medical Facility	0.003	0.002	0.002	-0.003	-0.004	-0.003
	(0.56)	(0.46)	(0.41)	(-0.70)	(-0.90)	(-0.92)
Rural Indicator	0.035	0.031	0.024	0.039	0.029	0.026
	(1.16)	(0.99)	(0.76)	(1.58)	(1.15)	(1.05)
Provincial Indicators	Yes	Yes	Yes	Yes	Yes	Yes
Observation	1479	1479	1479	1998	1998	1998
Pseudo R-squared	0.13	0.13	0.13	0.19	0.19	0.19

Note: Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 15: Probit and OLS Regressions Measuring the Effects of Income Inequality on Health Behaviors

	Probit	OLS	Probit	OLS
	Current Smoker	Cigarettes Per Day	Current Drinker	Drinking Frequency
	(1)	(2)	(4)	(5)
Community Gini	0.258***	6.662**	0.219**	0.970***
	(2.87)	(3.72)	(2.39)	(2.68)
Control Variables				
Income (1000 yuan)	0.013*	0.303**	0.030***	0.141***
	(1.67)	(2.06)	(3.80)	(4.74)
Education	-0.007***	-0.140***	0.006**	0.007
	(-2.64)	(-2.70)	(2.15)	(0.66)
Age	0.013***	0.461***	0.023***	0.106***
	(3.39)	(5.76)	(5.82)	(6.56)
Age Squared (/1000)	-0.182***	-5.498***	-0.261***	-1.088***
	(-4.49)	(-6.68)	(-6.34)	(-6.55)
Married Indicator	0.128***	2.811***	0.059**	0.395***
	(4.41)	(4.83)	(2.04)	(3.37)
Family Size	0.003	0.048	-0.002	0.003
	(0.53)	(0.39)	(-0.32)	(0.14)
Rural Indicator	0.008	0.694*	-0.005	0.022
	(0.39)	(1.70)	(-0.25)	(0.27)
Provincial Indicators	Yes	Yes	Yes	Yes
Observation	3172	3172	3172	3172
R-squared	0.03	0.06	0.03	0.04

Note: Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

Table 16: Probit Regressions Measuring the Effects of Relative Deprivation on Self-Reported Health Status

	Dependent Variable: Self-Reported Health Status (1=excellent or good, 0=fair or poor)			
	(1)	(2)	(3)	(4)
RDA (/1000)	0.003			
	(0.18)			
RDL		-0.004		
		(-0.30)		
RDI			-0.0001	
			(-0.08)	
Individual Rank				0.046*
				(1.91)
Control Variables				
Income (/1000)	0.035***	0.033***	0.034***	0.020*
	(3.69)	(3.10)	(3.70)	(1.75)
Income Squared	-0.002**	-0.002**	-0.002**	-0.001
	(-2.18)	(-1.96)	(-2.14)	(-1.18)
Education	0.002	0.002	0.002	0.003*
	(1.58)	(1.46)	(1.59)	(1.73)
Age	-0.011***	-0.011***	-0.011***	-0.011***
	(-4.62)	(-4.54)	(-4.61)	(-4.59)
Age Squared (/1000)	0.033	0.032	0.033	0.033
	(1.40)	(1.34)	(1.40)	(1.40)
Male Indicator	0.041***	0.041***	0.041***	0.040***
	(3.70)	(3.73)	(3.70)	(3.64)
Married Indicator	0.043***	0.041**	0.043***	0.043***
	(2.59)	(2.48)	(2.59)	(2.61)
Family Size	0.012***	0.013***	0.012***	0.012***
	(3.56)	(3.68)	(3.57)	(3.61)

Household Environment	0.130***	0.126***	0.130***	0.132***
	(8.94)	(8.64)	(8.97)	(9.03)
Distance to Medical Facility	-0.001	-0.001	-0.001	-0.001
	(-0.60)	(-0.71)	(-0.63)	(-0.76)
Rural Indicator	0.041***	0.040***	0.041***	0.036***
	(3.19)	(3.16)	(3.19)	(2.79)
Provincial Indicators	Yes	Yes	Yes	Yes
Observation	7300	7271	7298	7300
Pseudo R-squared	0.10	0.09	0.10	0.10

Note: Numbers in parentheses are t-statistics.

*, **, and *** represent significance levels of 10, 5, and 1 percent.

References:

Aghion, Philippe, Eve Caroli and Cecilia Garcia-Penalosa. 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature* 37, 1615-60.

Alesina, Alberto and Roberto Perotti. 1996. "Income Distribution, Political Instability, and Investment." *European Economic Review* 40, 1203-28.

Ben-Shlomo, Yoav, Ian R. White and Michael Marmot. 1996. "Does the Variation in Socio-Economic Characteristics of an Area Affect Mortality?" *British Medical Journal* 312(20), 1013-14.

Benjamin, Dwayne, Loren Brandt, and John Giles. 2003. "The Evolution of Income Inequality in Rural China." Working Paper, William Davidson Institute.

. 2004. "The Dynamics of Inequality and Growth in Rural China: Does Higher Inequality Impede Growth?" Working Paper, William Davidson Institute.

Benjamin, Dwayne, Loren Brandt, Paul Glewwe and Guo Li. 2002. "Markets, Human Capital and Income Inequality in an Economy in Transition: The Case of Rural China." in *Inequality Around the* World, Richard Freeman (ed.), Palgrave-MacMillan.

Blakely, Tony, A., Bruce P. Kennedy, and Ichiro Kawachi. 2001. "Socioeconomic Inequality in Voting Participation and Self-Rated Health." *American Journal of Public Health* 91(1), 99-104.

Blakely, Tony, A., Kimberley Lochner and Ichiro Kawachi. 2002. "Metropolitan Area Income Inequality and Self-Rated Health – A Multilevel Study." *Social Science and Medicine* 54, 65-77.

Bramall, Chris. 2001. "The Quality of China's Household Income Surveys." *China Quarterly* 167, 689-705.

Che, Jiahua and Yingyi Qian. 1998. "Institutional Environment, Community Government, and Corporate Governance: Understanding China's Township-Village Enterprises." *Journal of Law, Economics and Organization* 14, 1-23.

Chen, Hongyi and Scott Rozelle. 1999. "Leaders, Managers, and the Organization of Township and Village Enterprises in China." *Journal of Development Economics* 60, 529-57.

Cohen, Sheldon, Scott Line, Stephen Manuck, Bruce Rabin, Eugene Heise, and Jay Kaplan. 1997. "Chronic Social Stress, Social Status, and Susceptibility to Upper Respiratory Infections in Nonhuman Primates." *Psychosomatic Medicine* 59(3), 213-21.

Daly, Mary C., Greg J. Duncan, George A. Kaplan and John W. Lynch. 1998. "Macro-to-Micro Links in the Relation between Income Inequality and Mortality." *The Milbank Quarterly* 76(3), 315-39.

Deaton, Angus. 2001. "Relative Deprivation, Inequality, and Mortality." NBER Working Paper, #8099.

. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature* 41, 113-58.

Deaton, Angus and Christina Paxson. 1998. "Aging and Inequality in Income and Health." *American Economic Review* 88(2), 248-53.

Eibner, Christina E. and William N. Evans. 2001. "Relative Deprivation, Poor Health Habits and Mortality." Manuscript. College Park, Maryland: University of Maryland.

Feinstein, Jonathan S. 1993. "The Relationship between Socioeconomic Status and Health: A Review of the Literature." *The Milbank Quarterly* 71, 279-322.

Fiscella, Kevin and Peter Franks. 1997. "Poverty or Income Inequality as Predictor of Mortality: Longitudinal Cohort Study." *British Medical Journal* 314, 1724-28.

. 2000. "Individual Income, Income Inequality, Health and Mortality: What Are the Relationships?" *Health Services Research* 35(1), 307-18.

Gerdtham, Ulf-G. and Magnus Johannesson. 2004. "Absolute Income, Relative Income, Income Inequality, and Mortality." *Journal of Human Resources* 39(1), 228-48.

Gravelle, Hugh. 1998. "How Much of the Relation between Population Mortality and Unequal Distribution of Income is a Statistical Artefact?" British Medical Journal 316, 382-85.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy*, 80(2), 223-55.

Gustafsson, Bjorn and Shi Li. 2002. "Income Inequality within and across Counties in Rural China 1988 and 1995." *Journal of Development Economics* 69, 179-204.

Hsieh, Ching-Chi and Meredith D. Pugh. 1993. "Poverty, Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies." *Criminal Justice Review* 18, 182-202.

Idler, Ellen L. and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38, 21-37.

Kanbur, Ravi and Xiaobo Zhang. 1999. "Which Regional Inequality? The Evolution of

Rural-Urban and Inland-Coastal Inequality in China from 1983 to 1985." *Journal of Comparative Economics* 27, 686-701.

Kaplan, George A., Elsie R. Pamuk, John W. Lynch, Richard D. Cohen and Jennifer L. Balfour. 1996. "Inequality in Income and Mortality in the United States: Analysis of Mortality and Potential Pathways." *British Medical Journal* 312, 999-1003.

Kawachi, Ichiro and Bruce P. Kennedy. 1997. "The Relationship of Income Inequality to Mortality: Does the Choice of Indicator Matter?" *Social Science and Medicine* 45(7), 1121-27.

. 1999. "Income Inequality and Health: Pathways and Mechanisms." *Health Services Research* 34(1), 215-27.

Kawachi, Ichiro, Bruce P. Kennedy and Deborah Prothrow-Stith. 1997. "Social Capital, Income Inequality and Mortality." *American Journal of Public Health* 87(9), 1491-98.

Kawachi, Ichiro, Bruce P. Kennedy, Kimberly Lochner, et al. 1998. "Social Capital, Income Inequality and Firearm Violent Crimes." *Social Science and Medicine* 47(1), 7-17.

Kennedy, Bruce P., Ichiro Kawachi and Deborah Prothrow-Stith. 1996. "Income Distribution and Mortality: Cross-Sectional Ecological Study of the Robin Hood Index in the United States." *British Medical Journal* 312, 1004-07.

Kennedy, Bruce P., Ichiro Kawachi, Robert Glass and Deborah Prothrow-Stith. 1998. "Income Distribution, Socioeconomic Status, and Self-Rated Health in the United States." *British Medical Journal* 317, 917-21.

Khan, Azizur Rahman and Carl Riskin. 1998. "Income and Inequality in China: Composition, Distribution and Growth of Household Income, 1988 to 1995." *China Quarterly* 154, 221-53.

LeClere, Felicia B. and Mah-Jabeen Soobade. 2000. "The Effect of Income Inequality on the Health of Selected US Demographic Groups." *American Journal of Public Health* 90, 1892-97.

Li, Wei. 1997. "The Impact of Economic Reform on the Performance of Chinese State Enterprises, 1980-1989." *Journal of Political Economy* 105(5), 1080-1106.

Lin, Justin. 1992. "Rural Reforms and Agricultural Growth in China." *American Economic Review* 82(1), 34-51.

Lynch, John W., George A. Kaplan, Elsie R. Pamuk, et al. 1998. "Income Inequality and Mortality in Metropolitan Areas of the United States." *American Journal of Public Health* 88(7), 1074-80.

Lynch, John, George Davey Smith, Sam Harper, et al. 2004. "Is Income Inequality a Determinant of

Population Health? Part 1. A Systematic Review." *The Milbank Quarterly* 82(1), 5-99.

Marmot, Michael G. 1997. "Inequality, Deprivation and Alcohol Use." *Addiction* 92 (supp.1), s13.

Marmot, Michael G., Martin Shipley and Geoffrey Rose. 1984. "Inequalities in Death Specific Explanations of a General Pattern." *Lancet* 1(8384), 1003-1006.

Marmot, Michael G., George Davey-Smith, S. Stansfeld, et al. 1991. "Health Inequalities among British Civil Servants: The Whitehall II study." *Lancet* 337(8754), 1387-93.

Meara, Ellen. 1999. "Inequality and Infant Health." Manuscript. Boston, Massachusetts: Harvard Medical School.

Mellor, Jennifer M. and Jeffery Milyo. 2002. "Income Inequality and Health Status in the United States: Evidence from the Current Population Survey." *The Journal of Human Resources* 37, 510-39.

Morduch, Jonathan and Terry Sicular. 2002. "Rethinking Inequality Decomposition, with Evidence from Rural China." *Economic Journal* 112, 93-106.

Osler, Merete, Eva Prescott, Morten Gronbaek, et al. 2002. "Income Inequality, Individual Income, and Mortality in Danish Adults: Analysis of Pooled Data from Two Cohort Studies." *British Medical Journal* 324, 13-16.

Park, Albert and Scott Rozelle. 1998. "Reforming State-Market Relations in Rural China." *The Economics of Transition* 6(2), 461-80.

Preston, Samuel H. 1975. "The Changing Relation between Mortality and Level of Development." *Population Studies* 29, 231-48.

Qian, Yingyi and Barry Weingast. 1997. "Federalism as a Commitment to Reserving Market Incentives." *Journal of Economic Perspectives* 11(4), 83-92.

Robert, S. A. 1998. "Community-level Socioeconomic Status Effects on Adult Health." *Journal of Health and Social Behavior* 39, 18-37.

Rozelle, Scott. 1994. "Rural Industrialization and Increasing Inequality: Emerging Patterns in China's Reforming Economy." *Journal of Comparative Economics* 19, 362-91.

Runciman, W. G. 1966. *Relative Deprivation and Social Justice*. London: Routledge and Kegan Paul.

Shibuya, Kenji, Hideki Hashimoto and Eiji Yano. 2002. "Individual Income, Income Distribution, and Self-Rated Health in Japan: Cross-Sectional Analysis of Nationally Representative Sample." *British Medical Journal* 324, 16-19.

Shively, Carol A., Kathy Laber-Laird and Raymond F. Anton. 1997. "Behavior and Physiology of Social Stress and Depression in Female Cynomolgus Monkeys." *Biological Psychiatry* 41, 871-82.

Sicular, Terry. 1995. "Redefining State, Plan, and Market: China's Reforms in Agricultural Commerce." *China Quarterly* 144, 1020-46.

Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status." *Journal of Economic Perspectives* 13, 145-66.

Soobader, Mah-Jabeen and Felicia LeClere. 1999. "Aggregation and the Measurement of Income Inequality: Effects on Morbidity." *Social Science and Medicine* 48(6), 733-44.

Tsui, Kai-yuen. 1998. "Trends and Inequalities of Rural Welfare in China: Evidence from Rural Households in Guangdong and Sichuan." *Journal of Comparative Economics* 26, 783-804.

Waldmann, Robert J. 1992. "Income Distribution and Infant Mortality." *The Quarterly Journal of Economics* 107(4), 1283-302.

Wennemo, Irene. 1993. "Infant Mortality, Public Policy and Inequality - A Comparison of 18 Industrialized Countries, 1950-1985." *Sociology of Health and Illness* 15(4), 429-46.

Wilkinson, Richard G. 1992. "Income Distribution and Life-Expectancy." *British Medical Journal* 304, 165-68.

. 1996. *Unhealthy Societies: The Affliction of Inequality*. London, Routledge.

. 1997. "Health Inequalities: Relative or Absolute Material Standards?" *British Medical Journal* 314, 591-95.

Yitzhaki, Shlomo. 1979. "Relative Deprivation and the Gini Coefficient." *Quarterly Journal of Economics* 93(2), 321-24.

Zhao, Yaohui. 1999. "Leaving the Countryside: Rural to Urban Migration Decisions in China." *American Economic Review* 89(2), 281-86.